DYNAMIC SEGMENTATION OF BREAST TISSUE IN DIGITIZED MAMMOGRAMS

J. T. Neyhart[1], M. D. Ciocco[1], R. Polikar[1], S. Mandayam[1] and M. Tseng[2]
[1] Department of Electrical & Computer Engineering, Rowan University, Glassboro, NJ, USA
[2]Fox Chase Cancer Center, Philadelphia, PA, USA

Abstract-**The percentage of radiodense tissue in a mammogram has been used as a marker for determining breast cancer risk. In this paper, we present an image segmentation technique for identifying tissue and non-tissue regions of a digitized X-ray image. This procedure constitutes a vital step prior to subsequent processing for estimating the amount of radiodense tissue. The process involves the generation of a segmentation mask developed by using discrete wavelet transform techniques. Initial results have been promising, demonstrating the feasibility of the approach.**
Keywords - mammograms, segmentation, wavelets

I. INTRODUCTION

Mammography has emerged as a reliable technique for the early detection of breast cancer – the second leading cause of cancer-related mortality among American women [1]. The radiographic appearance of female breast consists of radiolucent (dark) regions due to fat and radiodense (light) regions due to connective and epithelial tissue. The amount of radiodense tissue can be used as a marker for predicting breast cancer risk. Women with radiodense tissue in more than 60-75% of the breast are at four to six times greater risk of developing breast cancer than those with lesser densities [2]. The estimation of radiodense tissue has traditionally been a subjective determination by trained radiologists, with very few published work describing quantitative measures [3, 4]. This paper presents results obtained at Rowan University during the course of a research project intended to support an investigation conducted at Fox Chase Cancer Center (FCCC) in Philadelphia, PA, examining the correlation between dietary patterns and breast density. We have developed image processing algorithms that automatically scan digitized mammogram images to locate the breast tissue region in the X-ray, segment the tissue into radiodense and radiolucent indications and quantify the amount and percentage of radiodense tissue.

The objective of the project described in this paper is to develop an image processing algorithm that dynamically segments a digitized mammogram image into tissue (breast) and non-tissue regions, prior to subsequent analysis for identifying the radiodense indications. This process involves the generation of a segmentation mask, as shown in Fig. 1. Research data has been obtained from an existing cohort of women enrolled in the Family Risk Analysis Program (FRAP) at FCCC.

This paper is organized as follows: a background of the research activity is presented in the introduction section, followed by the overall approach and a detailed description of the implementation procedure. Typical results are then presented along with comparative evaluations of the approach to the results of other techniques. Finally, conclusions drawn from this investigative study are presented.

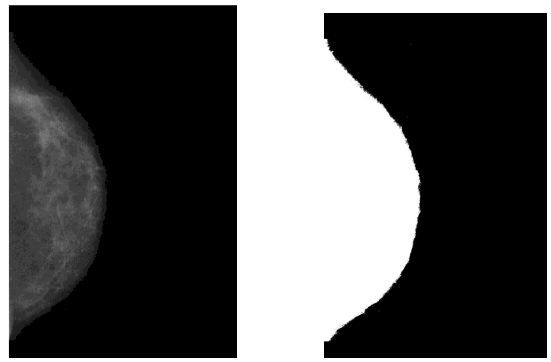


Fig. 1. Digitized mammogram image and its associated segmentation mask.

II. APPROACH

A block diagram describing the overall approach is shown in Fig. 2. The procedure described in this paper addresses the highlighted two blocks, namely, the mask generation and tissue segmentation.

The original mammogram X-rays were first digitized using a commercial film scanner. The images were then preprocessed to reduce the spatial resolution and improve the contrast using histogram adjustment. A segmentation mask was then designed to distinguish the tissue region from the film region. This mask is a template consisting of a binary matrix of size equal to that of the original image. The segmentation algorithm described below determines which sections of the image correspond to a tissue region, and assigns the value “1” (white) to the corresponding regions of the matrix. The rest of the matrix, corresponding to the non-tissue region, is set to “0” (black). This process allows us to subsequently identify radiodense regions in the image by concentrating on the tissue region only. However, determining an appropriate gray-level threshold for this conversion process is a non-trivial task. This is because the threshold cannot be an absolute value: it must respond to variations in signal intensity from image to image, but more importantly to local variations within the same image.

Identifying the radiodense tissue region in a segmented gray-level mammogram image essentially involves converting the 256 gray-level image to binary (black-and-white) format. Radiodense tissue pixels are assigned a gray-level value of 1 (white) whereas other pixels are assigned the value 0 (black). The dynamic density estimation based technique for detecting radiodense indications is described in [5].

Report Documentation Page

Report Date	Report Type	Dates Covered (from... to)
25OCT2001	N/A	-

Title and Subtitle Dynamic Segmentation of Breast Tissue in Digitized Mammograms	**Contract Number**
	Grant Number
	Program Element Number
Author(s)	**Project Number**
	Task Number
	Work Unit Number
Performing Organization Name(s) and Address(es) Department of Electrical & Computer Engineering, Rowan University, Glassboro, NJ	**Performing Organization Report Number**
Sponsoring/Monitoring Agency Name(s) and Address(es) US Army Research, Development & Standardization Group (UK) PSC 802 Box 15 FPO AE 09499-1500	**Sponsor/Monitor's Acronym(s)**
	Sponsor/Monitor's Report Number(s)
Distribution/Availability Statement Approved for public release, distribution unlimited	
Supplementary Notes Papers from the 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, October 25-28, 2001, held in Istanbul, Turkey. See also ADM001351 for entire conference on cd-rom., The original document contains color images.	
Abstract	
Subject Terms	
Report Classification unclassified	**Classification of this page** unclassified
Classification of Abstract unclassified	**Limitation of Abstract** UU
Number of Pages 4	

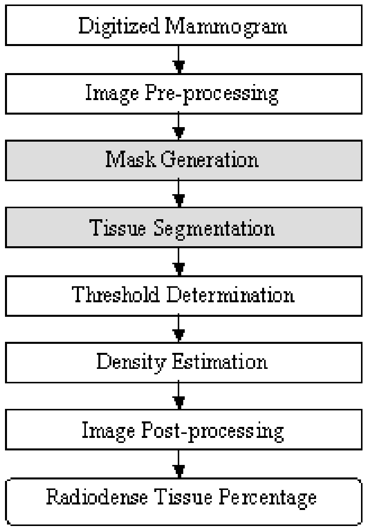


Fig. 2. Block diagram of research approach.

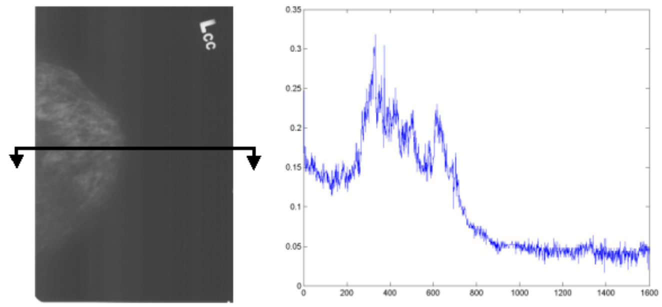
III. Procedure

A. Generation of the Segmentation Mask

The first step in the segmentation algorithm involves scanning the X-ray image to obtain the digitized mammogram. A line-scan of a single row in the digitized mammogram can be seen in Fig. 3. The goal is to identify the gray-level transition at the boundary of the tissue region. Note that this line-scan exhibits both local and global variations in gray-level. The global variation corresponds to the transition in the X-ray from the tissue to the non-tissue region, which is often obscured by the local variations corresponding to local changes in tissue density. The large variance of the local variations makes it impossible for a preset threshold to identify the tissue region (the left half of the scan). Furthermore, these local variations within the tissue region also make it very difficult to employ standard edge detection algorithms. We therefore use a discrete wavelet transform (DWT) based multiresolution decomposition [6] to simultaneously model both these variations in the gray-level for each line-scan in the original X-ray image. A polynomial fit is then used to correlate adjacent line scans to generate the final mask template. The mask template is a binary image consisting of white pixels corresponding to tissue regions and black pixels corresponding to non-tissue regions. The mask is then placed on the original mammogram image. The resulting segmented image contains the gray-level value of 0 (black) on all non-tissue regions, and the original gray-scale values in all tissue regions.

B. Discrete Wavelet Transform

The DWT is used for modeling local variations and global variations simultaneously in the line-scan for identifying the transition from the tissue to non-tissue region. In particular, the approximation coefficients of each line scan at a particular decomposition level were used to remove the local variations, while conserving the global variations.



(a) (b)

Fig. 3. (a) Original mammogram image and (b) line-scan

An extensive study of the image heuristics revealed that Daubechies mother wavelet with five vanishing moments provided the most optimal model. The fifth level approximation coefficients of the original signal provided the space-frequency information corresponding to the tissue boundary of the mammogram. This model is indicated by

$$\begin{aligned} Y_{id}(k) &= \sum_n Y_{(i-1)a}(n)\cdot g(2k-n) \\ Y_{ia}(k) &= \sum_n Y_{(i-1)a}(n)\cdot h(2k-n) \end{aligned} \qquad (1)$$

where Y_{ia} and Y_{id} are the approximation and detail (DWT) coefficients at the i^{th} level, respectively, and h(n) and g(n) are lowpass and highpass filters, obtained from Daubechies scaling and wavelet functions, respectively. At level zero, Y_{0a} represents the original raw line scan obtained from the image. Fig. 4 shows the associated wavelet decomposition tree for this system.

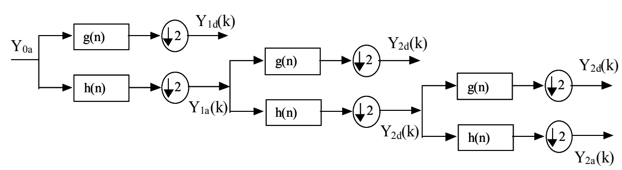


Fig. 4. Multiresolution wavelet decomposition tree

Fig. 5 illustrates a typical line scan and the corresponding 5th level approximation coefficients. From these coefficients, a threshold can be easily computed based on the statistical properties of the signal for identifying the tissue boundary.

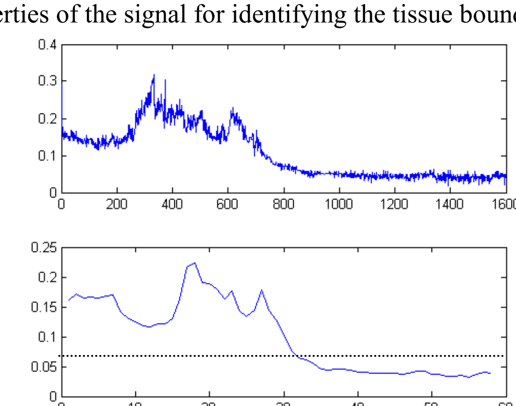
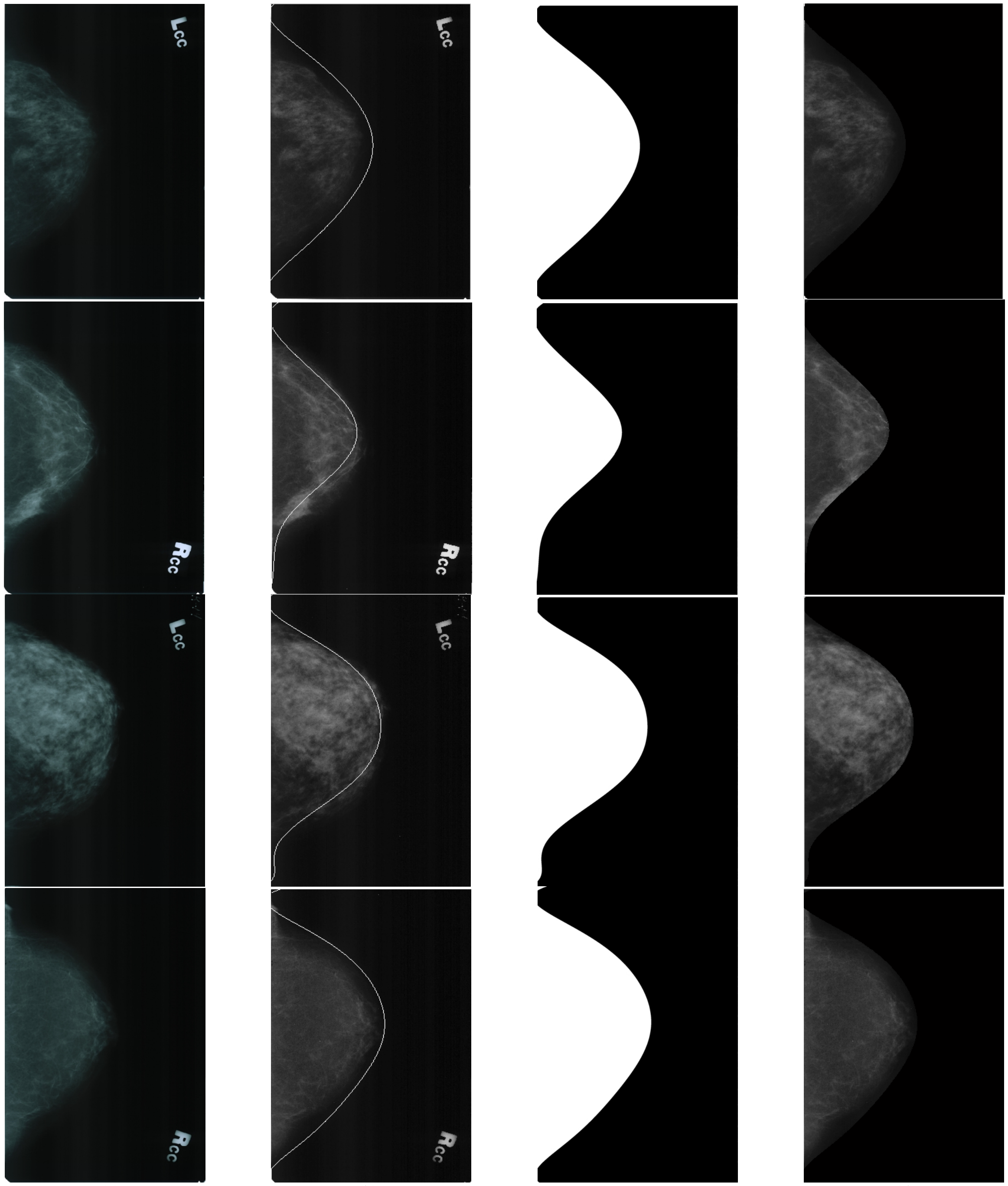


Fig.5. A typical line-scan and corresponding 5th level DWT approximation coefficients.



(a) (b) (c) (d)

Fig. 6. Four sets of data showing the original image, edge of tissue region as detected by using the DWT, the mask developed from this edge and the original image after multiplied by the mask

IV. RESULTS

Mammogram X-rays of left and right craniocaudal (lcc and rcc) as well as mediolateral-oblique views were obtained from a cohort of 30 women enrolled at FCCC's Family Risk Analysis Program (FRAP). The X-rays were digitized at 500 dpi using an Agfa medical-grade film scanner. Fig. 6 shows typical segmentation results obtained by applying the above-described approach. The images shown in Fig.6 indicate that the DWT is able to model the edge of the tissue region with sufficient accuracy. A quantitative analysis of the proximity of the mask contour to the tissue-film boundary can be made by comparing the dynamically generated mask to a segmentation mask that is generated manually. Table 1 shows the percentage difference and mean square errors (MSE) as calcu-

lated in (3) and (4), respectively, for each of the four cases analyzed in Fig. 5.

$$\%Diff = \left| \frac{(M_{dyn} - M_{man})}{M_{man}} \right| * 100 \qquad (2)$$

$$MSE = \frac{1}{N} \sum \left| \left(M_{man} - M_{dyn} \right)^2 \right| \qquad (3)$$

where M_{dyn} and M_{man} are the number of 1 (white) pixels in the DWT based segmentation mask and the manually generated mask, respectively, and N is the total number of pixels in the image.

TABLE I

Percentage and mean square error from manually developed reference mask

Patient ID	View	Hand Mask (pix)	Dynamic Mask (pix)	% diff	MSE ($x10^4$)
232217	lcc	829614	1049359	26.48	1.1641
235179	rcc	861665	653900	24.11	1.1507
245596	lcc	1197724	988142	17.49	1.2789
244231	rcc	1049230	1216404	15.93	0.7431

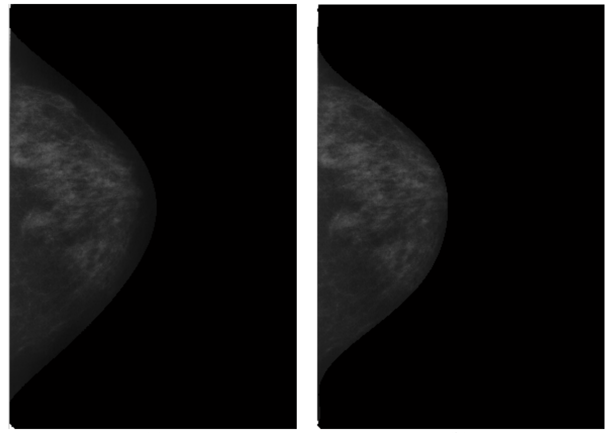
It should be noted that the manual segmentation provides the exact tissue boundary, and therefore all error figures are compared with respect to this manual segmentation benchmark. However, often times in image processing applications, the true performance of the algorithm can only be assessed by subjective visual evaluation of the resulting image. Furthermore, the MSE is only meaningful for relative comparison of different masks. Therefore, the numbers given in Table I should be interpreted within these guidelines.

V. Discussions and Conclusions

A DWT based approach for segmenting tissue regions in digitized mammogram X-rays is presented. The algorithm is capable of distinguishing local variations from global variations in tissue density, and hence able to identify the tissue to non-tissue boundary. The method is simple to implement and yields consistent results. It has been validated using data from an on-going study.

It can be argued that the DWT approximation of the line scans can also be obtained using a simple low pass filtering procedure, or by using any one of the multitude of edge detection algorithms available in image processing. However, it should be noted that the local variations within the tissue region renders edge detection extremely difficult.

Fig 7. compares segmentation results obtained by using the DWT based mask with a mask obtained by low pass filtering the individual line-scans. Although comparable results were obtained using a simple low pass filter based procedure, the DWT based approach provided us with a more robust estimation of the boundary.



(a) (b)

Fig. 7. Images segmented with
(a) DWT derived mask and (b) low pass filter derived mask

Furthermore, identifying the appropriate threshold was significantly simpler and computationally less expensive in the DWT based approach due to the 32-fold reduction in the number of pixels at the 5th level of the decomposition. Although computing the DWT at five levels is computationally more expensive then a single lowpass filtering, identifying the appropriate threshold takes considerably less amount of time in the DWT based approach, making the DWT based approach the faster overall algorithm.

Acknowledgments

We would like to thank Ms. Magdalena Kirlakovsky and Ms. Laura Coleman for digitizing the mammogram X-rays, Mr. David Chezem for setting up a secure online database for managing our data, Dr. Kathryn Evers, MD, for subjectively evaluating our algorithms, and the American Institute for Cancer Research for supporting this project.

References

[1] S. H. Landis, T. Murray, S. Bolden and P. A. Wingo, "Cancer statistics," CA Cancer Journal, Vol 48, pp. 6-29, 1998.

[2] N. F. Boyd, G. A. Lockwood, J. W. Byng, D. L. Tritchler and M. J. Yaffe, "Mammographic densities and breast cancer risk," *Cancer Epidemiology, Biomarkers and Prevention*, Vol. 7, pp. 1133-1144, December 1998.

[3] N. F. Boyd, J. W. Byng, R. A. Jong, E. K. Fishell, L. E. Little, A. B. Miller, G. A. Lockwood, D. L. Tritchler, M. J. Yaffe, "Quantitative classification of mammographic densities and breast cancer risk: results from the Canadian National Breast Screening Study," Journal of the National Cancer Institute, Vol. 87, No. 9, pp. 670-675, May 1995.

[4] M. J. Yaffe, N. F. Boyd, J. W. Byng, R. A. Jong, E. Fishell, G. A. Lockwood, L. E. Little, D. L. Tritchler, "Breast cancer risk and measured mammographic density," European Journal of Cancer Prevention, Vol 7 (suppl 1) pp. S47-S55, 1998.

[5] J. Neyhart, M. Kirlakovsky, S. Mandayam and M. Tseng, "Automated segmentation and quantitative characterization of radiodense tissue in digitized mammograms," to appear in *Proceedings of the 28th Annual Review of Progress in Quantitative NDE*, American Institute of Physics, New York, July 2001.

[6] G. Strang and T. Nguyen, *Wavelets and Filter Banks*, Wellesley-Cambridge Press, Wellesley, MA, 1996.